DOCUMENT RESUME

ED 290 756 SP 029 898

AUTHOR Haggard, Cynthia S.; Lang, Duaine C.
TITLE The Impact of Weighting Student Teacher Evaluation Variables.
PUB DATE 16 Feb 88
NOTE 11p.; Paper presented at the Annual Meeting of the Association of Teacher Educators (San Diego, CA, February 13-17, 1988).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Equated Scores; *Evaluation Criteria; Higher Education; *Interrater Reliability; *Measurement Techniques; Measures (Individuals); Preservice Teacher Education; *Student Teacher Evaluation; *Weighted Scores

ABSTRACT

The question of whether weighting (rank-ordering) of student teacher evaluation criteria, as determined by student and supervising teachers, university supervisors, and employing officials, results in better differentiation between and among student teachers is explored in this paper. Participants in the study were asked to rank order 15 evaluation items in terms of relative importance. The study included 193 student teachers, 194 supervising teachers, 27 university supervisors, and 59 employing officials. These multiple group rankings permitted an analysis of the question of whether different importance is attached to different items by different populations. Findings indicated that weighting items utilized in the evaluation of what usually appears to be a single exemplary population is not, as a rule, valid. (JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED290756.

THE IMPACT OF WEIGHTING STUDENT TEACHER EVALUATION VARIABLES

A Paper Presented at

The Association of Teacher Educators

Annual National Conference

San Diego, California

February 16, 1988

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. C. Lang

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Cynthia S. Haggard
Director of Student Teaching
Indiana University - South Bend, IN

and

Duaine C. Lang
Coordinator of Field Experiences
Indiana University - Bloomington, IN

SP 029 898

BEST COPY AVAILABLE

THE IMPACT OF WEIGHTING STUDENT TEACHER EVALUATION VARIABLES

Orwell's Animal Farm shocked us in the '40's with the pronouncement that "all animals are equal," but that "some animals are more equal." Today, those who work with student teachers can testify to the truth of the first statement -- they are all equally outstanding. The problem seems to be in trying to determine which "are more equal."

We seemingly all share the common dilemma of trying to devise an evaluation format that will be both fair (not influenced by undue subjectivity), and that which will truly discriminate.

Evaluations, whether by supervising teachers or university supervisors, in virtually all of our programs include some type of "check-list" rating of variables considered critical to instructional and teaching success, which are usually accompanied by a narrative statement. Analysis of the latter is hampered in that it typically only reinforces the check-list items, and does not lend itself readily to statistical treatment. Analysis of the former is usually of little value in that all items are treated as equivalent, and the scores are so universally inflated that everyone seems to share the same exalted degree of success.

This study posed the question whether weighting the items (ranking) and then comparing the item's weighted value with the evaluation given the student teacher for that item would result in a differentiated profile.

Student teachers, both pre and post student teaching, supervising teachers, university supervisors, and employing officials were requested to rank order the fifteen evaluation items. These multiple group rankings also permitted analysis of the question of whether different importance is attached to different items by different populations.

Analysis was also done in terms of the gender of the student teachers, the license level of both the student and supervising teachers, whether the student teaching was full or part semester, whether the student teaching was a single or dual assignment, the numbers of years of teaching experience and the number of student teachers previously supervised by the supervising teacher, and whether the university supervisor was a generalist or a specialist.

An additional question studied was whether individual evaluation items or combination of items served as predictors of success.

The study included 193 student teachers from IU-Bloomington during the second semester of the 1984-85 academic year. Forty-five of the student teachers had dual assignments. This resulted in a total of 238 item rankings and 238 final evaluations received. One hundred and ninty-four supervising teachers rank ordered items and responded to the questions relative to years of teaching experience and number of student teachers supervised. A total of twenty-seven university supervisors, seventeen generalists and ten specialists cooperated in ranking the items. And finally, a total of 59 employing officials, 89% of those surveyed, completed the item rankings.

Statistical treatments utilizing the Statistical Package for Social Studies (SPSS), dependent upon the variables being examined, included the ANOVA and Scheffé test, the t Test and Pooled Variance Estimate, Pearson Correlation, Frequencies, Crosstabs, and Breakdown. Significance levels for variance were set at both the .05 and .01 levels.

The fifteen items, taken from page one of the Final Evaluation Form, that the groups were asked to "rank" in terms of relative importance and that were also used as the basis for "evaluation" of the student teacher by the supervising teacher are shown as an Appendix to this paper.

The supervising teacher indicated on a twelve position Likert type scale an evaluation for each item. It was thus possible to assign a numeric value for each item, although the Form itself does not contain such values. This assigned numeric value could then be used for statistical comparison with the individual item's "rank-ordered" value.

The findings will be reported in terms of the questions posed as a part of this study.

Do Different Groups Rank Items the Same?

Analysis of the responses from the five groups, Pre Student Teaching Student Teachers, Post Student Teaching Student Teachers, Supervising Teachers, University Supervisors, and Employing Officials, produced significant differences for thirteen of the fifteen items. The difference for eleven of the thirteen was significant at the .01 level. The inter-group contributions to the level of significance, as revealed by the Scheffé Test, is of some interest. Additional information relative to this can be obtained from the authors.

The only two items for which no significant difference between the groups was found were Motivations and Enthusiasm and Relationship with Pupils.

Of perhaps even greater interest relative to the ranking of items was on the basis of other sub-group variables, as is shown in the following Table:

Table 1

Analysis Variables and Ranking Differences

Variable	Number of Significant Differences
Gender -- Student Teachers	3 of 15 items
License Level -- Student teachers Pre	4 of 15 items
License Level -- Student Teachers Post	5 of 15 items
License Level -- Supervising Teachers	4 of 15 items
Full or Part Semester Assignment Pre	1 of 15 items
Full or Part Semester Assignment Post	3 of 15 items
One or Two Student Teaching Assignments Pre	1 of 15 items
One or Two Student Teaching Assignments Post	1 of 15 items
Number of Student Teachers Supervised	1 of 15 items
Years of Teaching Experience	1 of 15 items
Type of University Supervisor	0 of 15 items

It would appear safe to assume that ranking of the evaluation items is significant only in terms of the actual groups themselves. Further analysis in terms of the group's sub-composition, at least in terms of the variables examined, was non productive.

Are Evaluations Received For Different Items Consistent?

The analyses relative to this question are quite in contrast to the item ranking issue. The answer to the basic question revealed that all student teachers on all items were consistently evaluated high -- the Combined Mean of all fifteen items was 10.749 on a 12.0 scale, and the range of means was only 1.02.

Sub-group analysis of the student teacher group revealed, however, that females were evaluated as significantly higher on 14 of the 16 items. (All fifteen items were summed to create a Combined Mean item.)

Elementary student teachers were evaluated as significantly better on ten of the sixteen items, and those who student taught for a full semester were adjudged as significantly better on nine of the sixteen items.

A rather interesting observation is that the three items for which the student teachers received the highest evaluation, (1) Health & Attendance, (2) Potential, and (3) Emotional Maturity, were the same items that they had rank ordered 13, 14, and 11 in terms of relative importance.

The same comparison of ranking importance vs evaluation received for Classroom Management produced a difference of 2 (high) to 14 (low), for Knowledge of Pupil Development of 7 to 13, and for Communications Skills of 3 to 9.

It would thus appear that those items student teachers valued most (ranked highest) were the ones upon which they were evaluated lowest, and what were valued least (ranked lowest) were evaluated the highest. The reader is reminded, however, that the range of ranking was from 1 to 15 whereas the evaluation range was only 1.02!

Will Item Ranking X Evaluation Received Produce A Difference?

The student teachers were divided into three groups in terms of how they had ranked each of the fifteen items. The High Group included those who had ranked an item as 1-5 in importance, the Middle Group ranked an item from 6-10, and the Low Group ranked the item from 11-15. The item's rank value was then multiplied by the actual evaluation value received.

Analysis of the relative position predicated upon the new values indicated no change in the composition of the three groups. This was particularly true

of the High and Low Groups. There was some internal movement within the Middle Group.

It would seem safe to observe that the "best" remain the best, and the "lowest" the lowest even though they rank-ordered the items differently. Again, without question, the primary reason can be attributed to the exceedingly small range of evaluation means.

Do Individual Items, or Combination of Items, Predict?

Pearson Product-Moment correlations were done for comparison of the evaluation value received for each item and the ranking value assigned each of the other items by the Student Teachers, both Pre and Post, and the Supervising Teachers. The process did not provide information that would change how student teachers are evaluated or that could define predictive variables or combinations of variables. The correlations were very weak, even when significant differences were found. An additional problem is that one cannot tell cause or sequence from correlations, so there is no way to determine which variable moved first or in which direction.

Summary

The answer to the question of whether weighting items utilized in the evaluation of student teachers can help differentiate between what usually appears to be a single exemplary population must be "No" on the basis of this study.

It was demonstrated that different groups do rank-order the items differently. The difference, however, is between the groups, and not attributable to differences within the groups; e.g. female vs male student teachers. This lack of difference does, however, provide some assurance that students on

different license programs, the number of years of teaching experience and the number of student teachers supervised by supervising teachers, and whether the university supervisor is a generalist or specialist, are not factors relative to the importance attributed to the different items.

It would appear that all student teachers receive consistently high evaluation on all items -- the total range of means was only 1.02. Examination of sub-group populations, however, revealed significant advantage for female vs male student teachers, for elementary as opposed to secondary or all grade student teachers, and for those who did a full semester of student teaching.

The basic question still remains. This study did provide some answers that can help in the design and implementation of student teaching and teacher preparation programs.

APPENDIX

INDIANA UNIVERSITY SCHOOL OF EDUCATION
FINAL EVALUATION OF STUDENT TEACHING/PRACTICUM

Educational Placement Office File Number _____

Name of Supervising Teacher _____

The family Educational Rights and Privacy Act of 1974 opens many student records for the student's inspection. The Law also permits the student to sign a waiver relinquishing such rights to inspect letters of recommendation. The applicant's signature below constitutes a waiver; no signature means the student will have the right to read this reference.

Date _____ Student's Signature _____

Name of Student _____ Date _____

Directions: Check the appropriate position on the continuum for each item.

The student should be evaluated as a beginning teacher.

INSTRUCTIONAL COMPETENCIES			
Subject Matter	Excellent knowledge	Adequate	Inadequate
Planning	Excellent daily & long-range planning	Adequate	Inadequate
Methods	Excellent selection, variety & implementation	Adequate	Inadequate
Instructional Materials	Excellent creativity and use	Adequate	Inadequate
Classroom Management	Excellent learning environment fostered	Adequate	Inadequate
PROFESSIONAL ATTRIBUTES			
Knowledge of Pupil Development	Excellent understanding	Adequate	Inadequate
Motivation and Enthusiasm	Highly motivated, enthusiastic toward teaching	Adequate	Inadequate
Relationship With Pupils	Excellent rapport and professional relationship	Adequate	Inadequate
Professional Commitment	Excellent commitment to improving quality of education	Adequate	Inadequate
Community Relations	Excellent relationship and understanding of parents and community	Adequate	Inadequate
PERSONAL ATTRIBUTES			
Communications Skills	Excellent speaking, writing, and listening skills	Adequate	Inadequate
Self-Evaluation Skills	Excellent ability to analyze own teaching related strengths and weaknesses	Adequate	Inadequate
Emotional Maturity	Excellent	Adequate	Inadequate
Health & Attendance	Excellent	Adequate	Inadequate
Potential	Excellent ability and commitment, will require minimal supervision	Adequate ability and commitment, will require moderate supervision	Questionable ability and/or commitment, will require extensive supervision

Signed _____
(supervising teacher)

Educational Placement Office Copy (white)
Office of Field Experiences Copy (yellow)
Student Copy (pink)
Supervising Teacher Copy (gold)